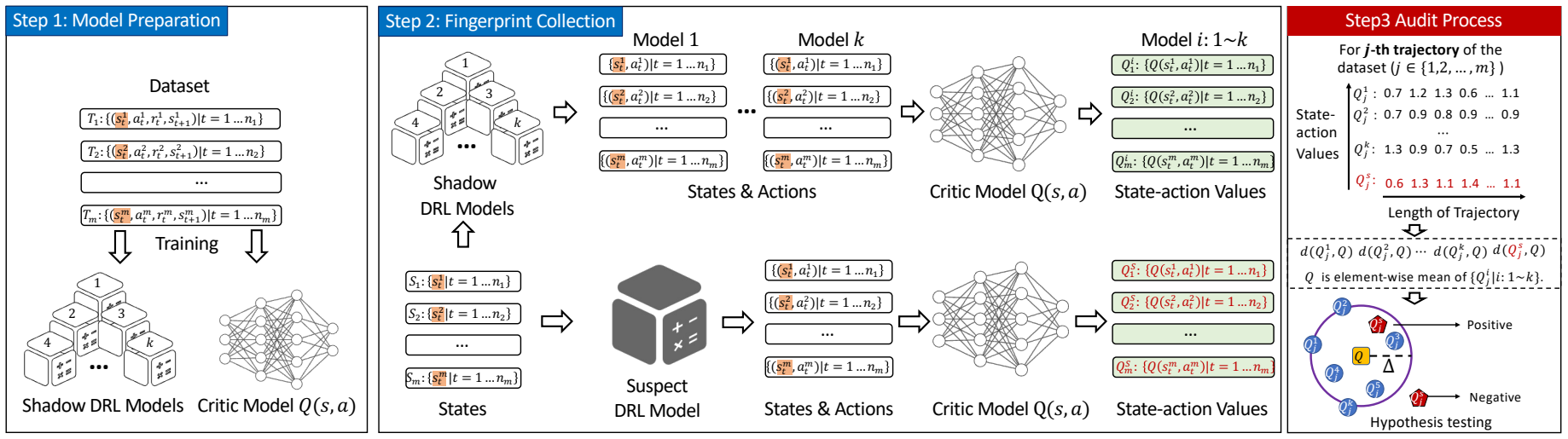
Step 1: Model Preparation
Dataset
Training
Shadow DRL Models
Critic Model $Q(s,a)$
Step 2: Fingerprint Collection
Model 1
Model k
Model i: 1~k
Shadow DRL Models
States & Actions
Critic Model Q(s,a)
State-action Values
States
Suspect DRL Model
States & Actions
Critic Model Q(s,a)
State-action Values
Step3 Audit Process
For j-th trajectory of the dataset ($j \in \{1,2,\dots,m\}$)
State-action Values
Length of Trajectory
$d(Q_j^1, Q)$ $d(Q_j^2, Q)$ $\cdots$ $d(Q_j^k, Q)$ $d(Q_j^s, Q)$
Q is element-wise mean of $\{Q_j^i | i: 1\sim k\}$.
Positive
Negative
Hypothesis testing